Exercises:

1. Let V be an inner product space with an orthonormal basis $\mathcal{B} := \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.
   (a) Prove that for any $\mathbf{x}, \mathbf{y} \in V$
   $$\langle \mathbf{x}, \mathbf{y} \rangle = \langle [\mathbf{x}]_{\mathcal{B}}, [\mathbf{y}]_{\mathcal{B}} \rangle,$$
   where the first inner product is from V and the second is from $\mathbb{F}^n$.
   (b) Prove **Parseval's identity**:
   $$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{v}_j \rangle \langle \mathbf{v}_j, \mathbf{y} \rangle \qquad \forall \mathbf{x}, \mathbf{y} \in V.$$

2. Let V be an inner product space and let P_E be the orthogonal projection onto a subspace $E \subset V$.
   (a) Prove that $P_E(\mathbf{v}) = \mathbf{v}$ if and only if $\mathbf{v} \in E$.
   (b) Prove that $P_E(\mathbf{w}) = \mathbf{0}$ if and only if $\mathbf{w} \perp E$.
   (c) Show that $P_E \circ P_E = P_E$ and $P_E \circ (I - P_E) = O$.

3. Let V be an inner product space and let $S \subset V$ be a subset. Define
   $$S^{\perp} := \{\mathbf{v} \in V : \mathbf{v} \perp \mathbf{x}\ \forall \mathbf{x} \in S\}.$$
   (a) Prove that $S^{\perp}$ is a subspace (even when S is not).
   (b) Show that $S \subset (S^{\perp})^{\perp}$.
   (c) Prove that $S = (S^{\perp})^{\perp}$ if and only if S is a subspace.
   (d) Prove that $(S^{\perp})^{\perp} = \operatorname{span} S$.

4. Let V be a finite-dimensional inner product space, let $E \subset V$ be a proper subspace ($E \neq V$ and $E \neq \{\mathbf{0}\}$), and let P_E be the orthogonal projection onto E.
   (a) Determine the spectrum $\sigma(P_E)$.
   (b) Prove that P_E is diagonalizable.
   (c) Find a diagonalization of P_E.

5. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^n$. Show that the system $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is orthonormal if and only if the matrix
   $$A = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$$
   is unitary.

6. Let $A \in M_{n\times n}$ be a normal matrix: $A^*A = AA^*$.
   (a) Show that an eigenvector $\mathbf{v}$ of A with eigenvalue λ is also an eigenvector of A^* with eigenvalue $\bar{\lambda}$.
   (b) Show that if λ, μ are **distinct** eigenvalues of A, then the eigenspaces $\operatorname{Ker}(A-\lambda I)$ and $\operatorname{Ker}(A-\mu I)$ are orthogonal.
   (c) Show that there exists a unitary matrix $U \in M_{n\times n}$ so that
   $$A = U \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} U^*,$$
   where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A counting multiplicities.

7*. (**Not collected**) Let V be a real vector space with norm $\|\cdot\|$. Assume the parallelogram identity holds:

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$$

for all $\mathbf{x}, \mathbf{y} \in V$. Define

$$\langle \mathbf{x}, \mathbf{y} \rangle := \frac{1}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{x}-\mathbf{y}\|^2.$$

In this exercise you will prove that the above map is an inner product.

(a) Show that $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ for all $\mathbf{x} \in V$. Use this to prove **non-negativity** and **non-degeneracy**.

(b) Prove **symmetry** directly; that is, show that $\langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in V$.

(c) Show that $\langle 2\mathbf{x}, \mathbf{y} \rangle = 2\langle \mathbf{x}, \mathbf{y} \rangle$ and $\langle \frac{1}{2}\mathbf{x}, \mathbf{y} \rangle = \frac{1}{2}\langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in V$.

(d) Show that $\langle \mathbf{w}+\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{w}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{w}, \mathbf{x}, \mathbf{y} \in V$.
[**Hint:** use $\mathbf{w}+\mathbf{x} \pm \mathbf{y} = \mathbf{w} \pm \frac{1}{2}\mathbf{y} + \mathbf{x} \pm \frac{1}{2}\mathbf{y}$.]

(e) Show that $\langle -\mathbf{x}, \mathbf{y} \rangle = -\langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in V$.

(f) Show that $\langle n\mathbf{x}, \mathbf{y} \rangle = n\langle \mathbf{x}, \mathbf{y} \rangle$ and $\langle \frac{1}{n}\mathbf{x}, \mathbf{y} \rangle = \frac{1}{n}\langle \mathbf{x}, \mathbf{y} \rangle$ for all $n \in \mathbb{Z}$ and all $\mathbf{x}, \mathbf{y} \in V$.

(g) Show that $\langle q\mathbf{x}, \mathbf{y} \rangle = q\langle \mathbf{x}, \mathbf{y} \rangle$ for all $q \in \mathbb{Q}$ and all $\mathbf{x}, \mathbf{y} \in V$.

(h) Use the previous parts to prove **rational linearity**; that is,

$$\langle p\mathbf{w} + q\mathbf{x}, \mathbf{y} \rangle = p\langle \mathbf{w}, \mathbf{y} \rangle + q\langle \mathbf{x}, \mathbf{y} \rangle,$$

for all $p, q \in \mathbb{Q}$ and all $\mathbf{w}, \mathbf{x}, \mathbf{y} \in V$.

(i) Use the parallelogram identity and the triangle inequality to show $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$.

(j) Show that

$$|\langle \alpha\mathbf{x}, \mathbf{y} \rangle - \alpha\langle \mathbf{x}, \mathbf{y} \rangle| \leq 2|\alpha - q|\|\mathbf{x}\|\|\mathbf{y}\|$$

for any $\alpha \in \mathbb{R}$, $q \in \mathbb{Q}$, and $\mathbf{x}, \mathbf{y} \in V$.

(j) Use the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$ (that is, any real number has a sequence of rational numbers converging to it) to prove $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$ for all $\alpha \in \mathbb{R}$ and all $\mathbf{x}, \mathbf{y} \in V$.

(k) Use the previous parts to prove **linearity**.

Solutions:

1. (a) We know there are scalars $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ such that

$$[\mathbf{x}]_{\mathcal{B}} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \qquad \text{and} \qquad [\mathbf{y}]_{\mathcal{B}} = \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n.$$

This implies

$$\mathbf{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Now, using linearity and conjugate linearity we have

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n, \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n \rangle = \sum_{i,j=1}^{n} \alpha_i\bar{\beta}_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle$$

Recalling that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal system, this reduces to

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \alpha_i\bar{\beta}_i\|\mathbf{v}_i\|^2 = \sum_{i=1}^{n} \alpha_i\bar{\beta}_i = \langle [\mathbf{x}]_{\mathcal{B}}, [\mathbf{y}]_{\mathcal{B}} \rangle.$$

□

(b) Letting the α_i's and β_i's be as in the previous part, we have for each $j = 1, \ldots, n$ that

$$\langle \mathbf{x}, \mathbf{v}_j \rangle = \langle \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n, \mathbf{v}_j \rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \alpha_j \|\mathbf{v}_j\|^2 = \alpha_j.$$

By a similar computation, we obtain $\langle \mathbf{v}_j, \mathbf{y} \rangle = \bar{\beta}_j$. Thus by part (a) we obtain

$$\sum_{j=1}^{n} \langle \mathbf{x}, \mathbf{v}_j \rangle \langle \mathbf{v}_j, \mathbf{y} \rangle = \sum_{j=1}^{n} \alpha_j \bar{\beta}_j = \langle [\mathbf{x}]_{\mathcal{B}}, [\mathbf{y}]_{\mathcal{B}} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$$

□

2. (a) ($\Longrightarrow$): Suppose $P_E(\mathbf{v}) = \mathbf{v}$. Then by definition of the orthogonal projection $\mathbf{v} = P_E(\mathbf{v}) \in E$.
($\Longleftarrow$): Suppose $\mathbf{v} \in E$. Then recall that by a theorem from lecture we have

$$\|\mathbf{v} - P_E(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{x}\| \qquad \forall \mathbf{x} \in E.$$

In particular, this holds for $\mathbf{x} = \mathbf{v}$ which makes the right-hand side zero. This forces the left-hand side to be zero and so we must have $\mathbf{v} - P_E(\mathbf{v}) = \mathbf{0}$, or $\mathbf{v} = P_E(\mathbf{v})$. □

(b) ($\Longrightarrow$): Suppose $P_E(\mathbf{w}) = \mathbf{0}$. Observe that $\mathbf{w} = \mathbf{w} - \mathbf{0} = \mathbf{w} - P_E(\mathbf{w})$, which is orthogonal to E by definition of the orthogonal projection.
($\Longleftarrow$): Suppose $\mathbf{w} \perp E$. Let $\mathbf{v} \in E$, then

$$\langle P_E(\mathbf{w}), \mathbf{v} \rangle = \langle P_E(\mathbf{w}), \mathbf{v} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle = \langle P_E(\mathbf{w}) - \mathbf{w}, \mathbf{v} \rangle = -\langle \mathbf{w} - P_E(\mathbf{w}), \mathbf{v} \rangle = 0,$$

where the last equality holds since $(\mathbf{w} - P_E(\mathbf{w})) \perp E$ by definition of the orthogonal projection. The above holds, in particular, when $\mathbf{v} = P_E(\mathbf{w})$, which yields $\langle P_E(\mathbf{w}), P_E(\mathbf{w}) \rangle = 0$ and so $P_E(\mathbf{w}) = \mathbf{0}$ by non-degeneracy. □

(c) Let $\mathbf{v} \in V$. Then $P_E(\mathbf{v}) \in E$ by part (a). Thus, by part (a) again we have $P_E(P_E(\mathbf{v})) = P_E(\mathbf{v})$. That is, $P_E \circ P_E(\mathbf{v}) = P_E(\mathbf{v})$. Since $\mathbf{v} \in V$ was arbitrary, this implies $P_E \circ P_E = P_E$.
For the second equation, simply observe

$$P_E \circ (I - P_E) = P_E \circ I - P_E \circ P_E = P_E - P_E = O,$$

where we have used the first part. □

3. (a) Let $\mathbf{v}, \mathbf{w} \in S^\perp$ and let α, β be scalars. Then for any $\mathbf{x} \in S$ we have by linearity of the inner product that

$$\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{x} \rangle = \alpha \langle \mathbf{v}, \mathbf{x} \rangle + \beta \langle \mathbf{w}, \mathbf{x} \rangle = \alpha 0 + \beta 0 = 0.$$

Thus $\alpha\mathbf{v} + \beta\mathbf{w} \in S^\perp$ and so $S^\perp$ is closed under addition and scalar multiplication. Also, $\langle \mathbf{0}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in S$, and so $\mathbf{0} \in S^\perp$. Hence S is a subspace. □

(b) Let $\mathbf{x} \in S$. Then for any $\mathbf{v} \in S^\perp$ we have $\langle \mathbf{x}, \mathbf{v} \rangle = 0$ so that $\mathbf{x} \perp \mathbf{v}$. Since $\mathbf{v} \in S^\perp$ was arbitrary, we have $\mathbf{x} \in (S^\perp)^\perp$. □

(c) ($\Longrightarrow$): Suppose $S = (S^\perp)^\perp$. Then letting $S_1 := S^\perp$, we have $S = S_1^\perp$ and so S is a subspace by part (a).
($\Longleftarrow$): Suppose S is a subspace. From part (b) we know $S \subset (S^\perp)^\perp$, and so it suffices to show $S \supset (S^\perp)^\perp$. Let $\mathbf{y} \in (S^\perp)^\perp$ and let P_S be the orthogonal projection onto S. Let $\mathbf{w} := \mathbf{y} - P_S(\mathbf{y})$, which is orthogonal to S by definition of the orthogonal projection. That is, $\mathbf{w} \in S^\perp$ and so $P_S(\mathbf{y}) \perp \mathbf{w}$, but also $\mathbf{y} \perp \mathbf{w}$ since $\mathbf{y} \in (S^\perp)^\perp$. Thus

$$\langle \mathbf{w}, \mathbf{w} \rangle = \langle \mathbf{y} - P_S(\mathbf{y}), \mathbf{w} \rangle = \langle \mathbf{y}, \mathbf{w} \rangle - \langle P_S(\mathbf{y}), \mathbf{w} \rangle = 0 - 0 = 0.$$

By non-degeneracy we must have $\mathbf{w} = \mathbf{0}$, which implies $\mathbf{y} = P_S(\mathbf{y})$. Then by Exercise 2.(a), we have $\mathbf{y} \in S$. □

(d) Let $E := \operatorname{span} S$, which is a subspace. Now, the inclusion $S \subset E$ implies $E^\perp \subset S^\perp$. Indeed, if $\mathbf{w} \in E^\perp$ then $\mathbf{w} \perp \mathbf{v}$ for all $\mathbf{v} \in E$. In particular, $\mathbf{w} \perp \mathbf{x}$ for all $\mathbf{x} \in S \subset E$ and thus $\mathbf{w} \in S^\perp$. Thus $E^\perp \subset S^\perp$, and by the same reasoning we have $(S^\perp)^\perp \subset (E^\perp)^\perp$. By part (c), we then know the latter set is E and so $(S^\perp)^\perp \subset E$. On the other hand, by parts (a) and (b), $(S^\perp)^\perp$ is a subspace containing S. It must therefore contain the span of S; that is, $E \subset (S^\perp)^\perp$. Hence $E = (S^\perp)^\perp$, as claimed. □

4. (a) Let $\lambda \in \sigma(P_E)$ and let $\mathbf{v}$ be an associated eigenvector. Since $P_E \circ P_E = P_E$ by Exercise 2.(c), we have

$$\lambda \mathbf{v} = P_E(\mathbf{v}) = P_E \circ P_E(\mathbf{v}) = P_E(\lambda \mathbf{v}) = \lambda^2 \mathbf{v},$$

which implies $(\lambda - \lambda^2)\mathbf{v} = \mathbf{0}$. Since $\mathbf{v} \neq \mathbf{0}$ (by virtue of being an eigenvector), we must have $\lambda - \lambda^2 = 0$. Noting that $\lambda - \lambda^2 = \lambda(1-\lambda)$, we see that this is only possible if $\lambda \in \{0, 1\}$. Hence $\sigma(P_E) \subset \{0,1\}$. On the other hand, since E is a propert subspace there exists non-zero vectors $\mathbf{v} \in E$ and $\mathbf{w} \in E^\perp$ and by Exercise 2.(a),(b) we have

$$P_E(\mathbf{v}) = \mathbf{v} = 1\mathbf{v}$$
$$P_E(\mathbf{w}) = \mathbf{0} = 0\mathbf{w},$$

so that $1, 0 \in \sigma(P_E)$. Thus $\sigma(P_E) = \{0, 1\}$. □

(b) Let $\mathcal{B}$ be a basis for E and let $\mathcal{C}$ be a basis for $E^\perp$. From the proof of Corollary 5.16 which says

$$\dim(E) + \dim(E^\perp) = \dim(V),$$

we know that $\mathcal{B} \cup \mathcal{C}$ is a basis for V. By the previous part, we know this is a basis of eigenvectors for P_E and therefore P_E is diagonalizable. □

(c) Let $\mathcal{B}$ and $\mathcal{C}$ be as in the previous part. Then

$$[P_E]_{\mathcal{B}\cup\mathcal{C}}^{\mathcal{B}\cup\mathcal{C}} = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix},$$

where the number of 1's along the diagonal is $\dim(E)$. □

5. Note that

$$A^* = \begin{pmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_n^* \end{pmatrix}.$$

It follows that

$$(A^*A)_{i,j} = \mathbf{v}_i^* \mathbf{v}_j = \langle \mathbf{v}_i, \mathbf{v}_j \rangle.$$

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal system, we obtain

$$(A^*A)_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} = (I_n)_{i,j}.$$

Thus $A^*A = I_n$ which implies A^* is a left-inverse of A. Since A is square, it is **the** inverse of A. That is, $A^* = A^{-1}$ and so A is unitary.

Conversely, suppose A is unitary. Then by our computation above we have

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = (A^*A)_{i,j} = (I_n)_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Hence $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal system. □

6. (a) We must show $A^*\mathbf{v} = \bar{\lambda}\mathbf{v}$, and we will use the non-degeneracy of the inner product to accomplish this:

$$\begin{aligned} \left\langle A^*\mathbf{v} - \bar{\lambda}\mathbf{v}, A^*\mathbf{v} - \bar{\lambda}\mathbf{v}\right\rangle &= \langle A^*\mathbf{v}, A^*\mathbf{v}\rangle - \lambda\langle A^*\mathbf{v}, \mathbf{v}\rangle - \bar{\lambda}\langle \mathbf{v}, A^*\mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle \\ \text{(Proposition 5.6)} \quad &= \langle \mathbf{v}, AA^*\mathbf{v}\rangle - \lambda\langle \mathbf{v}, A\mathbf{v}\rangle - \bar{\lambda}\langle A\mathbf{v}, \mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle \\ (A \text{ is normal}) \quad &= \langle \mathbf{v}, A^*A\mathbf{v}\rangle - \lambda\langle \mathbf{v}, \lambda\mathbf{v}\rangle - \bar{\lambda}\langle \lambda\mathbf{v}, \mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle \\ &= \langle A\mathbf{v}, A\mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle - |\lambda^2|\langle \mathbf{v}, \mathbf{v}\rangle + |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle \\ &= \langle \lambda\mathbf{v}, \lambda\mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle \\ &= |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle - |\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle = 0 \end{aligned}$$

Hence $A^*\mathbf{v} - \bar{\lambda}\mathbf{v} = \mathbf{0}$, or equivalently $A^*\mathbf{v} = \bar{\lambda}\mathbf{v}$. □

(b) Let $\mathbf{v} \in \mathrm{Ker}(A - \lambda I)$ and $\mathbf{w} \in \mathrm{Ker}(A - \mu I)$. Since $\lambda \neq \mu$, one of them must be non-zero. Without loss of generality we may assume $\lambda \neq 0$. Using part (a) and Proposition 5.6 we have

$$\langle \mathbf{v}, \mathbf{w}\rangle = \frac{\lambda}{\lambda}\langle \mathbf{v}, \mathbf{w}\rangle = \frac{1}{\lambda}\langle \lambda\mathbf{v}, \mathbf{w}\rangle = \frac{1}{\lambda}\langle A\mathbf{v}, \mathbf{w}\rangle = \frac{1}{\lambda}\langle \mathbf{v}, A^*\mathbf{w}\rangle = \frac{1}{\lambda}\langle \mathbf{v}, \bar{\mu}\mathbf{w}\rangle = \frac{\mu}{\lambda}\langle \lambda, \mu\rangle.$$

Thus

$$\left(1 - \frac{\mu}{\lambda}\right)\langle \mathbf{v}, \mathbf{w}\rangle = 0.$$

Since $\lambda \neq \mu$, we know the first factor above is non-zero. Thus the second factor, $\langle \mathbf{v}, \mathbf{w}\rangle$, must be zero. That is, $\mathbf{v} \perp \mathbf{w}$. Since $\mathbf{v} \in \mathrm{Ker}(A - \lambda I)$ and $\mathbf{w} \in \mathrm{Ker}(A - \mu I)$ were arbitrary, the eigenspaces are orthogonal. □

(c) Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of A. From lecture we know that normal matrices are diagonalizable, so it follows that

$$\sum_{k=1}^{r} \dim(\mathrm{Ker}(A - \lambda_k I)) = n.$$

For each $k = 1, \ldots, r$, let $\mathcal{B}_k$ be an orthonormal basis for the eigenspace $\mathrm{Ker}(A - \lambda_k I)$. Note that $\mathcal{B}_k$ consists of eigenvectors of A with eignevalue λ_k. By the previous part, $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ is an orthonormal system and hence linearly independent. The above equation implies $\mathcal{B}$ contains n elements and hence is an orthonormal basis for $\mathbb{F}^n$ (consisting of eigenvectors for A). Let $U \in M_{n\times n}$ be the matrix whose columns are the vectors in $\mathcal{B}$. Then

$$A = U \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} U^{-1}$$

By Exercise 5, U is a unitary matrix and so $U^{-1} = U^*$. □

7*. (a) We have

$$\langle \mathbf{x}, \mathbf{x}\rangle = \frac{1}{4}\|\mathbf{x} + \mathbf{x}\|^2 - \frac{1}{4}\|\mathbf{x} - \mathbf{x}\|^2 = \frac{1}{4}\|2\mathbf{x}\|^2 - \frac{1}{4}\|\mathbf{0}\|^2 = \frac{1}{4}4\|x\|^2 - 0 = \|x\|^2.$$

Thus $\langle \mathbf{x}, \mathbf{x}\rangle = \|x\|^2 \geq 0$ for all $\mathbf{x} \in V$, and $\langle \mathbf{x}, \mathbf{x}\rangle = 0 \Leftrightarrow \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$, which implies $\langle \cdot\,, \cdot \rangle$ is non-negative and non-degenerate. □

(b) We compute

$$\langle \mathbf{y}, \mathbf{x}\rangle = \frac{1}{4}\|\mathbf{y}+\mathbf{x}\|^2 - \frac{1}{4}\|\mathbf{y}-\mathbf{x}\|^2 = \frac{1}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}\|-(\mathbf{x}-\mathbf{y})\|^2 = \frac{1}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}|-1|^2\|\mathbf{x}-\mathbf{y}\|^2 = \langle \mathbf{x}, \mathbf{y}\rangle.$$

□

(c) We compute

$$\langle 2\mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\|2\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}\|2\mathbf{x}-\mathbf{y}\|^2 = \frac{1}{4}\|(\mathbf{x}+\mathbf{y})+\mathbf{x}\|^2 - \frac{1}{4}\|(\mathbf{x}-\mathbf{y})+\mathbf{x}\|^2. \tag{1}$$

Now, using the parallelogram identity, we have

$$\begin{aligned}\|(\mathbf{x}\pm\mathbf{y})+\mathbf{x}\|^2 &= -\|(\mathbf{x}\pm\mathbf{y})-\mathbf{x}\|^2 + 2\|\mathbf{x}\pm\mathbf{y}\|^2 + 2\|\mathbf{x}\|^2 \\ &= -\|\pm\mathbf{y}\|^2 + 2\|\mathbf{x}\pm\mathbf{y}\|^2 + 2\|\mathbf{x}\|^2 \\ &= -\|\mathbf{y}\|^2 + 2\|\mathbf{x}\pm\mathbf{y}\|^2 + 2\|\mathbf{x}\|^2\end{aligned}$$

Substituting this into Equation (1) gives

$$\begin{aligned}\langle 2\mathbf{x}, \mathbf{y}\rangle &= \frac{1}{4}(-\|\mathbf{y}\|^2 + 2\|\mathbf{x}+\mathbf{y}\|^2 + 2\|\mathbf{x}\|^2) - \frac{1}{4}(-\|\mathbf{y}\|^2 + 2\|\mathbf{x}-\mathbf{y}\|^2 + 2\|\mathbf{x}\|^2) \\ &= \frac{2}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{2}{4}\|\mathbf{x}-\mathbf{y}\|^2 = 2\langle \mathbf{x}, \mathbf{y}\rangle.\end{aligned}$$

From this it follows that $\langle \mathbf{x}, \mathbf{y}\rangle = \left\langle 2\frac{1}{2}\mathbf{x}, \mathbf{y}\right\rangle = 2\left\langle \frac{1}{2}\mathbf{x}, \mathbf{y}\right\rangle$. Thus $\left\langle \frac{1}{2}\mathbf{x}, \mathbf{y}\right\rangle = \frac{1}{2}\langle \mathbf{x}, \mathbf{y}\rangle$. □

(d) We compute

$$\begin{aligned}\langle \mathbf{w}+\mathbf{x}, \mathbf{y}\rangle &= \frac{1}{4}\|\mathbf{w}+\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{w}+\mathbf{x}-\mathbf{y}\|^2 \\ &= \frac{1}{4}\|(\mathbf{w}+\frac{1}{2}\mathbf{y})+(\mathbf{x}+\frac{1}{2}\mathbf{y})\|^2 - \frac{1}{4}\|(\mathbf{w}-\frac{1}{2}\mathbf{y})+(\mathbf{x}-\frac{1}{2}\mathbf{y})\|^2\end{aligned} \tag{2}$$

Now, using the parallelogram identity, we have

$$\begin{aligned}\|(\mathbf{w}\pm\frac{1}{2}\mathbf{y})+(\mathbf{x}\pm\frac{1}{2}\mathbf{y})\|^2 &= -\|(\mathbf{w}\pm\frac{1}{2}\mathbf{y})-(\mathbf{x}\pm\frac{1}{2}\mathbf{y})\|^2 + 2\|\mathbf{w}\pm\frac{1}{2}\mathbf{y}\|^2 + 2\|\mathbf{x}\pm\frac{1}{2}\mathbf{y}\|^2 \\ &= -\|\mathbf{w}-\mathbf{x}\|^2 + 2\|\mathbf{w}\pm\frac{1}{2}\mathbf{y}\|^2 + 2\|\mathbf{x}\pm\frac{1}{2}\mathbf{y}\|^2\end{aligned}$$

Substituting this into Equation (2) gives

$$\begin{aligned}\langle \mathbf{w}+\mathbf{x}, \mathbf{y}\rangle =& \frac{1}{4}\left(-\|\mathbf{w}-\mathbf{x}\|^2 + 2\|\mathbf{w}+\frac{1}{2}\mathbf{y}\|^2 + 2\|\mathbf{x}+\frac{1}{2}\mathbf{y}\|^2\right) \\ &- \frac{1}{4}\left(-\|\mathbf{w}-\mathbf{x}\|^2 + 2\|\mathbf{w}-\frac{1}{2}\mathbf{y}\|^2 + 2\|\mathbf{x}-\frac{1}{2}\mathbf{y}\|^2\right) \\ =& \frac{2}{4}\|\mathbf{w}+\frac{1}{2}\mathbf{y}\|^2 - \frac{2}{4}\|\mathbf{w}-\frac{1}{2}\mathbf{y}\|^2 + \frac{2}{4}\|\mathbf{x}+\frac{1}{2}\mathbf{y}\|^2 - \frac{2}{4}\|\mathbf{x}-\frac{1}{2}\mathbf{y}\|^2 \\ =& 2\left\langle \mathbf{w}, \frac{1}{2}\mathbf{y}\right\rangle + 2\left\langle \mathbf{x}, \frac{1}{2}\mathbf{y}\right\rangle \\ =& 2\left\langle \frac{1}{2}\mathbf{y}, \mathbf{w}\right\rangle + 2\left\langle \frac{1}{2}\mathbf{y}, \mathbf{x}\right\rangle \\ =& \langle \mathbf{y}, \mathbf{w}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle \\ =& \langle \mathbf{w}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle.\end{aligned}$$

□

(e) First observe that

$$\langle \mathbf{0}, \mathbf{y}\rangle = \frac{1}{4}\|\mathbf{0}+\mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{0}-\mathbf{y}\|^2 = \frac{1}{4}\|\mathbf{y}\|^2 - \frac{1}{4}\|-\mathbf{y}\|^2 = 0.$$

Thus by the previous part we have

$$\langle \mathbf{x}, \mathbf{y}\rangle + \langle -\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}+(-\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{0}, \mathbf{y}\rangle = 0,$$

so that $\langle -\mathbf{x}, \mathbf{y}\rangle = -\langle \mathbf{x}, \mathbf{y}\rangle$. □

(f) We will first prove $\langle n\mathbf{x}, \mathbf{y}\rangle = n\langle \mathbf{x}, \mathbf{y}\rangle$ for $n \in \mathbb{N}$ by induction on n. The base case of $n = 1$ is immediate. So suppose it holds for n. Then we have by part (d) that

$$\langle (n+1)\mathbf{x}, \mathbf{y}\rangle = \langle n\mathbf{x} + \mathbf{x}, \mathbf{y}\rangle = \langle n\mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle = n\langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle = (n+1)\langle \mathbf{x}, \mathbf{y}\rangle .$$

So by induction we have the claimed formula. Now, for $-n$, we simply apply the above and part (e). Finally, observe that

$$\langle \mathbf{x}, \mathbf{y}\rangle = \left\langle n\frac{1}{n}\mathbf{x}, \mathbf{y}\right\rangle = n\left\langle \frac{1}{n}\mathbf{x}, \mathbf{y}\right\rangle,$$

so that $\frac{1}{n}\langle \mathbf{x}, \mathbf{y}\rangle = \left\langle \frac{1}{n}\mathbf{x}, \mathbf{y}\right\rangle$. □

(g) For any $q \in \mathbb{Q}$, we can write $q = \frac{n}{m}$ for $n \in \mathbb{Z}$ and $m \in \mathbb{N}$. So using part (f) twice we get

$$\langle q\mathbf{x}, \mathbf{y}\rangle = n\left\langle \frac{1}{m}\mathbf{x}, \mathbf{y}\right\rangle = n\frac{1}{m}\langle \mathbf{x}, \mathbf{y}\rangle = q\langle \mathbf{x}, \mathbf{y}\rangle .$$

□

(h) Using part (d) and (g) we have

$$\langle p\mathbf{w} + q\mathbf{x}, \mathbf{y}\rangle = \langle p\mathbf{w}, \mathbf{y}\rangle + \langle q\mathbf{x}, \mathbf{y}\rangle = p\langle \mathbf{w}, \mathbf{y}\rangle + q\langle \mathbf{x}, \mathbf{y}\rangle .$$

□

(i) First note that by the parallelogram identity that

$$\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\|\mathbf{x}+\mathbf{y}\|^2 - \frac{1}{4}(2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 - \|\mathbf{x}+\mathbf{y}\|^2) = \frac{1}{2}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right).$$

We also have

$$-\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\|\mathbf{x}-\mathbf{y}\|^2 - \frac{1}{4}(2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{2}\left(\|\mathbf{x}-\mathbf{y}\|^2 - 2\|\mathbf{x}\|^2 - 2\|\mathbf{y}\|^2\right).$$

Now, the triangle inequality implies

$$\|\mathbf{x}\pm\mathbf{y}\|^2 \leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 = \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2.$$

Substituting this into our earlier computations yields

$$\pm\langle \mathbf{x}, \mathbf{y}\rangle \leq \frac{1}{2}(2\|\mathbf{x}\|\|\mathbf{y}\|) = \|\mathbf{x}\|\|\mathbf{y}\|.$$

Thus $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$. □

(j) Let $\alpha \in \mathbb{R}$ and let $q \in \mathbb{Q}$. Using parts (d) and (g) we have

$$\langle \alpha\mathbf{x}, \mathbf{y}\rangle = \langle (\alpha - q + q)\mathbf{x}, \mathbf{y}\rangle = \langle (\alpha - q)\mathbf{x}, \mathbf{y}\rangle + \langle q\mathbf{x}, \mathbf{y}\rangle = \langle (\alpha - q)\mathbf{x}, \mathbf{y}\rangle + q\langle \mathbf{x}, \mathbf{y}\rangle .$$

Thus

$$\langle \alpha\mathbf{x}, \mathbf{y}\rangle - \alpha\langle \mathbf{x}, \mathbf{y}\rangle = \langle (\alpha - q)\mathbf{x}, \mathbf{y}\rangle + (q-\alpha)\langle \mathbf{x}, \mathbf{y}\rangle .$$

Then using part (i) we get

$$|\langle \alpha\mathbf{x}, \mathbf{y}\rangle - \alpha\langle \mathbf{x}, \mathbf{y}\rangle| \leq |\langle (\alpha - q)\mathbf{x}, \mathbf{y}\rangle| + |q-\alpha||\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|(\alpha-q)\mathbf{x}\|\|\mathbf{y}\| + |\alpha-q|\|\mathbf{x}\|\|\mathbf{y}\| \leq 2|\alpha-q|\|\mathbf{x}\|\|\mathbf{y}\|,$$

where we have used homogeneity of the norm in the last step. □

(k) Let $\alpha \in \mathbb{R}$. Since $\mathbb{Q}$ is dense there is a sequence $(q_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} q_n = \alpha.$$

In particular,

$$\lim_{n\to\infty} |\alpha - q_n| = |\alpha - \alpha| = 0.$$

So using part (j) we see that

$$|\langle \alpha\mathbf{x}, \mathbf{y}\rangle - \alpha\langle \mathbf{x}, \mathbf{y}\rangle| = \lim_{n\to\infty} |\langle \alpha\mathbf{x}, \mathbf{y}\rangle - \alpha\langle \mathbf{x}, \mathbf{y}\rangle| \leq \lim_{n\to\infty} 2|\alpha - q_n|\|\mathbf{x}\|\|\mathbf{y}\| = 0.$$

Thus $\langle \alpha\mathbf{x}, \mathbf{y}\rangle - \alpha\langle \mathbf{x}, \mathbf{y}\rangle = 0$ or $\langle \alpha\mathbf{x}, \mathbf{y}\rangle = \alpha\langle \mathbf{x}, \mathbf{y}\rangle$. □

(l) Using parts (d) and (k) we have for any $\alpha, \beta \in \mathbb{R}$ that

$$\langle \alpha \mathbf{w} + \beta \mathbf{x}, \mathbf{y} \rangle = \langle \alpha \mathbf{w}, \mathbf{y} \rangle + \langle \beta \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{w}, \mathbf{y} \rangle + \beta \langle \mathbf{x}, \mathbf{y} \rangle .$$

□